DOCUMENT RESUME

ED 051 263

TM 000 537

AUTHOR Cunningham, Donald J.
TITLE Formative Evaluation of Replicable Forms of Instruction.
PUB DATE Feb 71
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Academic Achievement, Classification, Data Analysis, *Data Collection, Decision Making, Educational Environment, *Evaluation Methods, *Formative Evaluation, *Instructional Materials, Measurement Goals, *Program Improvement, Student Characteristics, Summative Evaluation, Teacher Characteristics, Teaching Methods

ABSTRACT

Current thinking and research on the nature of formative evaluation of replicable forms of instruction are summarized. Methodologies appropriate for other evaluation roles (e.g. summative) are shown to be often inappropriate for the formative role. This paper attempts to classify sources of information concerning the merit of educational materials and to suggest appropriate methodologies of formative evaluation. The data sources are of three broad types: external sources, internal sources, and contextual sources. Several types of data under each heading are examined to illustrate the classification system. (Author)

ED051263

FORMATIVE EVALUATION OF REPLICABLE FORMS OF INSTRUCTION[1]

Donald J. Cunningham

Indiana University

Current thinking and research on the nature of formative evaluation of replicable forms of instruction are summarized. Methodologies appropriate for other evaluation roles (e.g. summative) are shown to be often inappropriate for the formative role. This paper attempts to classify sources of information concerning the merit of educational materials and to suggest appropriate methodologies of formative evaluation. The data sources are of three broad types: external sources, internal sources, and contextual sources. Several types of data under each heading are examined to illustrate the classification system.

TM 000 537

[1] Paper presented at the annual meeting of the American Educational Research Association, New York, New York, February, 1971.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

FORMATIVE EVALUATION OF REPLICABLE FORMS OF INSTRUCTION

Donald J. Cunningham
Indiana University

This paper will focus on the evaluation, revision, and improvement of replicable forms of instruction. By replicable is meant those instructional materials which can be repeated in all their essential details and so includes such things as textbooks, films, workbooks, programmed sequences, videotapes, audiotapes, slides, and the like. The most obvious exclusion from consideration here is the classroom teacher although through the use of videotape or film it would be possible to replicate the teacher's performance. By evaluation is meant the collection and processing of informationfor purposes of making decisions about these materials. The terms revisions and improvements refer to a concern in this presentation for a particular type of decision, a decision typically made by the developers of instructional materials. The validation of instructional materials has been of central concern to the field of instructional development but a· yet there seems to be neither a general theory of evaluation to guide in the selection of information that is most relevant and appropriate nor any generally accepted methodology to indicate how information should be collected, processed, and interpreted. It is hoped that this presentation will help clarify some theoretical and methodological issues in evaluation and contribute to the further development of both the theory and practice of evaluation.

Evaluation may be defined as the collection and use of information to make decisions about certain entities. In the field of education

these entities could be such things as a curriculum, textbook, slide projector, teacher, school system, and the like. Decisions concerning such entities are made every day and it is the purpose of evaluation to provide information which will assist decision makers. This general definition tells us little about the field of evaluation, however, and, in fact, there is presently considerable debate as to the proper nature and function of educational evaluation. For example, the tradutional identifacation of evaluation with the one process of constructing and adminintering paper and pencil tests to individual students has been roundly criticized (Cronbach, 1963; Hastings, 1966). Others have rejected the notion that evaluation is a form of and therefore constrained by the reauirements of judgmental data as well as description data in educational evaluation is being expanded to include a wide variety of decisions. In other words, the scope of evaluation has been expanded to include all of the possible roles that it may serve.

But with the proliferation of possible roles of evaluation comes a danger that the essential goal of evaluation will be obscured. Scriven (1967) has emphasized this distinction between the goal of evaluation and the role of evaluation and it will be useful to summarize his comments here. The goal of evaluation is to answer questions concerning the merit of an entity, questions such as "How well does this entity work?" "Does it work better than other such entities?" "Is this entity worth the cost of developing it for widespread use?" Merit may be considered in both an absolute or comparitive sense, but in either case, the goal of

evaluation is the location of merit. The role of evaluation, on the other hand, can take many forms and refers to the use to which the evaluation will be put, to the type of decision that will be made. Thus an evaluation may be used to develop and improve a film, to support a decision on whether to adopt or reject a particular textbook, to determine the need for a new reading program, to support the existence of the athletic program, and so forth. As the number of possible roles that an evaluation can serve increases, the likelihood that particular theoretical considerations and methodological strategies will apply equally well to all evaluation roles decreases. Many writers in the field of evaluation seem to have assumed that it will be possible to develop a theory of evaluation which will encompass all of these roles and have devoted considerable attention to the developement of such a theory. The utility of this approach will not be debated here. However an emphasis on the central unity of the various evaluation roles may require quite different kinds of information and quite different methodologies for dealing with that information. A failure to appreciate this point has led to the use of inappropriate tools and techniques in many evaluations.

For instance, suppose we wished to evaluate textbook A. What sort of information should we collect? If the role of our evaluation is summative, then one very appropriate type of information may be a comparison of the score on a standarized achievement test of students using textbook A and students using another textbook, say textbook B. This information would be very valuable in deciding which of these textbooks to adopt in a school system. If, however, the role of the evaluation is to provide data that will be of use in revising and improving textbook A,

then something more than traditional achievement test scores is needed. For one thing, standardized tests are usually designed to measure individual differences in student achievement rather than the extent to which a particular textbook has achieved its objective. Thus, as Glaser (1963) points out, criterion-referenced, not norm-referenced achievement tests are needed in such evaluations. Criterion-referenced tests will be much more useful in locating those points in the textbook that need revision. A comparison of textbook A with textbook B might be completely irrelevant to an evaluation for purposes of revision.

Two roles which have recieved considerable attention in recent years are those which Scriven (1967) has labeled formative and summative evaluation. Formative evaluation is concerned with instrument development and revision; that is, with the collection of information which is fed back to the developer of the instrument (e,g. textbook author, lesson writer, film maker, etc.) with the intent that on the basis of these data the instrument will be revised and improved (if necessary, of course). Summative evaluation, according to Scriven, is considered with the evaluation of information concerning an educational instrument for purpose of making judgments about the instrument as a whole. Summative evaluation is therefore concerned with finished instruments, instruments in which no further refinements are planned. As defined by Scriven, however, the distinction between these two roles is not entirely satisfactory for it is difficult to concieve of many situations in education where developement ceases, where only summative, after the fact data would be collected. Stake (1969) has improved these definitions by proposing that formative and summative evaluation roles can be distinguished in terms of the audience of the evaluation. Formative evaluation refers to the gathering of information which would be of use to the <u>developers</u> of instructional materials, those who are trying to choose or produce the parts, the elements which will combine to form a

successful whole. Summative evaluation would refer to the collection of imformation which would be of use to the consumer of the instructional materials, those who seek to determine what they are getting for their money, what they can expect from this product, whether this product is better than others, and so forth. Information of use to developers will sometimes but not necessarily overlap with information of use to consumers and vice versa. Particular aspects of the materials which are evaluated and particular methodologies used to assess the merit of the materials, may, and usually do differ in formative evaluation and summative evaluation.

The fact that the appropriateness of a methodologies of evaluation depends to a large extent upon the role of evaluation does not imply that there is one correct methodology of formative evaluation. Scriven (1967) points out that the goal of evaluation is the assessment of merit. But where is merit to be found? What sources of information are relevant? Programmed instruction enthusiasts have have made a useful distinction between internal and external sources of information (Lumsdaine, 1965). Internal characteristics are those which can be determined from an inspection of the materials themselves (e.g. content coverage, readability, relative emphasis of certain topics, etc.). External characteristics refer to the effects of the materials on the behavior of students, teachers, parents, and so forth.[2] This distinction is similar to that made by Scriven (1967) between pay-off and instrinsic evaluations. Instrinsic and pay-off evaluations are procedures or methods of conducting evaluation and should not be confused with evaluation roles. An intrinsic evaluation can be used in either a formative or summative role, for instance. An intrinsic

---

[2] The term external characteristics, as defined by Lumsdaine (1965), refers to some other things than just the effects of the materials (e.g. history of the program developement, reputation of the author, and so forth. In this paperm however, the term external characteristics will refer only to the external effects of the materials.

evaluation would be concerned with the internal characteristics of the instructional materials while a pay-off evaluation would concentrate on the external effects of the materials. Scriven's example is worth repeating here:

> "If you want to evaluate a tool, say an ax, you might study the design of the bit, the weight distribution, the steel alloy used, the grade of hickory in the handle, etc., or you might study the kind and speed of cuts it makes in the hands of a good ax-man . . . . The first approach involves an appraisal of the instrument itself . . . (while) the second approach proceeds via an examination of the effects of the . . . . instrument." (Scriven, 1967, page 53).

Of course, there are probably no examples of either a pure intrinsic or pay-off evaluation but it is quite obvious that different instructional developers place different emphasis or priority on these two methods of evaluation. For some, the principle concern is with the "elegance" of the instructional materials, or with the breadth of coverage, or with the structure of the subject matter being taught. For others, the major (only?) concern is with changes in student behavior. Both schools of instructional development may engage in formative evaluation but certainly the types of information each will collect will be different, or, at least, information will be weighted differently. It is not the intention here to judge whether one orientation or the other is "better" or more "useful". In my opinion, the merit in instructional materials is to be found in both internal and external characteristics and information concerning both these characteristics are essential to sound curriculum development.

But internal and external information are not sufficient. Information concerning the conditions under which the materials are expected to function is also needed, information such as the type of students and teachers for which the materials are intended, the availability of various instructional aides, characteristics of school settings, and so forth.

Developers of instructional materials make frequent assumptions concerning the context within which their material will function but seldom gather information to test the adequacy of their assumptions. Contextual information, as I shall call it, is also very important for the formative evaluation of instructional materials.

Information concerning products which might possibly be of use in the formative evaluation of these products thus fall into three main classes: internal, external, and contextual. Each of the categories can be further subdivided as shown in Table 1. No pretense is made that the classification system and examples presented in Table 1 are exhaustive of the sorts of information that a formative evaluator might (not should) collect. Nevertheless this classification should serve to remind evaluators who have concentrated their efforts on the collection of certain types of information of the possibilities for providing a more complete evaluation of their product.

Internal Characteristics

Internal information has been defined as information obtainable from an inspection of the materials themselves. For purposes of this discussion, the term "materials" will include not only those materials which are actually presented to the student, but also other materials which often accompany instructional materials. These additional materials include such things as a list of the objectives of the instructional materials, the preface of a textbook, directions to the teacher on using the materials, a bibliography indicating the source of the program content, a biography of the author, test items which the author feels measure desired outcomes of the program, and so forth. These supplementary materials should be included in a determination of internal information. It should be noted that the availability of such supplementary materials will influence

greatly the type of evaluation which is conducted. For instance, if the objectives of the instructional materials are listed by the author, the evaluation of those objectives will be of a different sort than if the objectives are not stated but must be inferred from an inspection of the materials. In the first case the information is primarily descriptive while in the second the information is primarily judgmental.

Two major categories of internal information are proposed: descriptive information and critical appraisal. Descriptive information refers to the generally non-controversial and objective information which can be generated by inspecting the product. Critical appraisals are also made by inspecting the materials but this information is much more value-laden, subjective, and controversial. This is not to imply that descriptive information is free from subjectivity and value, however. The selection of information to describe and instruments to measure are decisions heavily laden with value. Nevertheless, I believe a meaningful distinction can be made between these two types of information. Descriptive information, while not generally considered to be a source of information concerning revisions, will, when linked with other types of information, indicate where revisions might be made.

Descriptive information. Table 1 lists four examples of descriptive information. Physical specifications refer to descriptions of the product itself including media characteristics. For instance, a programmed textbook can be described in terms of frame types, step size, blackout ratio, confirmation procedure, type of print, type of paper, frequency of illustrations, nature of student response, location of review frames, location of terminal frames, frequency and type of prompts, and so forth. One can also note the availability of supplementary materials such as panels, criterion tests, teacher guides, outside readings, and the like. It is obvious from the partial listing that many of these features are peculiar

to programmed media and might not be appropriate descriptors of another medium such as film. Obviously other physical characteristics would be described for a film or textbook or videotape. Edmonson et al. (1931) and Cronbach (1955) are useful sources for the determination of the physical characteristics of textbooks. Markle (1969) and Klaus (1965) suggest many characteristics of programmed materials which should be inventoried in a description of internal characteristics. Spottiswoode (1958) and Kracauer (1960) should serve the same purpose in the film medium.

The Rationale of instructional materials refers to the author's beliefs about the nature of the educative process and the relationship between his materials and that process. Stake (1967) believes that all instructional materials have a rationale even if it is only implicit. Statements of rationale are often found in textbook prefaces or can be inferred from the general objectives of the program. If a statement of rationale is not available from the author, then the evaluator may want to state what he feels is the rationale for the materials based upon his inspection of the materials. It should be noted however, that such a procedure goes beyond mere description.

Goals and Objectives refer to the more particular intents that the author has in mind for his materials. Goals and objectives can be classified in a number of ways which might prove useful to the formative evaluator: level of specifity (general, specific, behavioral), domain represented (cognitive, affective, psychomotor), process vs product, form (verbal statement vs sample test items), and so forth. Once again, if the objectives of the instructional materials are not listed, the evaluator might wish to infer these objectives by inspecting the materials. Stake (1967), Krathwohl (1964, 1965), Mager (1962), and Kibler et al. (1970) are useful sources of information concerning the description of rationales,

goals, and objectives.

Content information refers to the substance and form of presentation of the subject matter content. For instance, one might describe the subject matter coverage through the use of a topical outline or content analysis of the materials. Information concerning the sources of the subject matter content is often located in prefaces, bibliographies, credits, acknowledgements, and so forth. The relative emphasis of topics can be revealed by certain types of content analysis. Sequence can often be characterized as hierarchical, linear, spiral, random, independent, etc. Selected references which should assist in the description of content are Berelson (1954), Kerlinger (1964), Gagne (1970), and Easley et al. (1967).

The fourth example of descriptive information listed in Table 1 is Pedagogical strategy. Often it is possible to characterize instructional materials as to the theory of learning or instruction upon which they seemed to be based. Often an author will state such a bias in the preface to his book, while at other times his bias is obvious from an inspection of his materials. For instance, it might be possible to categorize materials as teacher-centered or student-centered depending on the extent to which the student is responsible for choosing and directing the use of the materials. I know of no references which deal expressly with the characterization of instructional materials as to their pedagogical strategy but some of the techniques of content analysis cited above and of classroom observation (Rosenshine, 1970) might be adapted for this purpose.

Critical Appraisals. Like descriptive information, critical appraisals are made by inspecting the materials themselves. Unlike descriptive information, however, critical appraisals are, by their very nature, subjective since human judgment plays a central role. Several

classes of people whose opinions ought to be considered are listed in Table 1. Methods for collecting critical appraisals and other judgment data have been discussed by Stake (1970). Critical appraisals have played a rather large part in the evaluation of certain forms of instruction. Hoban (1943) conducted a classic study in the evaluation of films in which critical appraisals by teachers and students were the primary data source. The use of critical appraisals in the formative evaluation of instructional materials has been severely criticized in some quarters and probably justifiably so. In many cases critical appraisal has been the only information collected concerning the instructional materials. Such an evaluation is likely to be inadequate just as would an evaluation which included only external information. One type of information is this category which has received very little attention is that supplied by the author himself. Author hindsight, that is, the author taking a second look at his materials is undoubtedly the most important source of information concerning revisions, particularly in the early stages of the development of a product. No lesson writer or film maker produces a product with which he is satisfied on the first draft or on the first take. A process not unlike that observed in the Skinner Box and called successive approximation seems to take place. To my knowledge, very little attention has been paid to the author as evaluator of his own instructional materials. Some preliminary studies conducted at the University of Illinois by Easley (1967) have dealt with this process but only in a limited way. Some interesting questions which come to mind on this point are, for instance:

1. Do authors differ as to the types of revisions they perform at various stages in the development of materials?

2. Do the author's attitudes toward internal and external information influence the sorts of decisions he makes?

3. Can the critical-appraisal capabilities of authors be modified through exposure to various theories of instruction?

It should be noted that other systems for the appraisal of internal characteristics of instructional materials have been developed and should be considered as possible alternatives to the one listed here. Stevens and Morrissett (1968) list six major types of information which one should collect concerning instructional materials: descriptive characteristics, rationale and objectives, antecedent conditions, content, instructional theory and teaching strategies, and overall judgments. Tyler (1970) suggests seven categories of information: rationale, specifications, appropriateness, effectiveness, conditions, practicality, and disseminations. Eash (1969) has developed an instrument for appraising curriculum materials around four constructs: statement of objectives, organization of the materials, methodology, and evaluation. All of these systems seem to have been developed to assist the consumer of instructional materials make decisions about adoption. They are, therefore, only suggestive of the types of information which might be useful to the developer of instructional materials.

External Information

External information refers to information collected by observing the effects of the materials on the behaviors of students, teachers, parents, and so forth. Types of and methodologies for the collection of external information have received a great deal more attention in the evaluation literature, hence will only be briefly reviewed here.

A paper by Metfessel and Michael (1967) has provided a useful categorization of external information, which, in modified form is listed in Table 1. Metfessel and Michael (1967) also suggest many techniques for the measurement of these behaviors. Sjogren (1970), Rosenshine (1970), Westbury (1970), and Stake (1970) review various techniques and method-

ologies which might assist in the collection of external information.

One methodology which has not been dealt with in these sources but which holds great promise for the formative evaluation of instructional materials is what I have chosen to call oral problem solving (after Bloom and Broder, 1950). Essentially this technique consists of placing the author (or his agent) with students as they use the materials. At first only one or two students are used and these students are asked to "think aloud" as they work through the materials. The authors often make assumptions concerning the mental processes students will employ when using the materials but seldom check to see if these assumptions are correct. Students can also help the author locate ambiguities, errors of sequence, and so forth. This procedure is, of course, the one which has been widely employed in the development of programmed materials (see Markle, 1970). There is no need, however, to limit its use to programmed materials. Textbook authors could profit greatly by having students read and comment on earlier drafts of texts. Film makers could have students react to the story board of the film to uncover any problems which might exist before the film is shot. Research on the usefulness of this methodology is lacking so as yet only common sense suggestions can be offered (see Scott and Yelon, 1969; Markle, 1970). It is hoped that considerably more attention will be devoted to this potentially very useful methodology.

Contextual Information

Contextual information is not usually collected in formative evaluations but is of great importance if the developer is to produce a product with some generalizability. If the materials are not used in the contexts for which they were designed, then the author may have to

revise some of his objectives or instruction. Often authors assume that their materials will be effective in a wide variety of contexts yet only infrequently do they tryout their products in different contexts. Table 1 lists several categories of contextual information which might be considered in a formative evaluation. The students for whom the materials are intended can be characterized in terms of experiential background, intelligence, interest, socio-economic status, and the like. Teachers will vary on such characteristics as previous experience, teaching style, personality, interest, and so forth. The curricular context refers to the type(s) of concurrent course work, availability of instructional aids, etc. Such information could have a profound effect on the predicted effectiveness of the instructional materials. In addition to these contextual variables, the type of school and the community can affect the performance of the materials. A community or school which is open to innovation might achieve different results with the materials than a school or community which is rather hostile to innovation (see, for instance, Anderson, 1969). Other variables which might be considered are parental attitudes, home environment, and so forth.

Many of the methods and techniques suitable for the collection of external information are also suitable for contextual information. Sjogren (1970) has reviewed some additional techniques for assessing school and community characteristics. Stufflebeam (1968) has urged the collection of such information in what he calls "context evaluation". It is likely that evaluators will have to draw upon the methodologies of sociology, economics, anthropology, and so forth to assist in the full description of context.

Table 1

A CLASSIFICATION OF INFORMATION OF POSSIBLE USE

IN FORMATIVE EVALUATIONS

I Internal Information

A. Descriptive information.

1. Physical specifications
2. Rationale, goals, and objectives
3. Content
4. Pedagogical strategy
5. Other

B. Critical appraisal.

1. Author (developer)
2. Experts (subject matter, media, psychologists, etc.)
3. Students using the materials
4. Teachers using the materials
5. Relevant others

II External Information

A. Assessment of the effects of the materials on student behavior.

1. achievement
2. attitude
3. skill
4. interest
5. committment
6. other

B. Assessment of the effects of the materials on teacher behavior.

1. attitude
2. interest
3. committment
4. competency
5. teaching strategy
6. other

C. Assessment of the effects of the materials on the behavior of relevant others.

1. parents
2. administrators
3. teachers not using the materials
4. students not using the materials
5. the community
6. other

Table 1
continued

III Contextual Information

A. Student characteristics

B. Teacher characteristics

C. School characteristics

D. Community characteristics

E. Curricular context

F. Other relevant elements in the learning environment

Discussion

It has been argued in this paper that the information needs of a formative evaluation are likely to be different from those of a summative evaluation. It has also been argued that the information likely to be of use in a formative evaluation falls into three major classes: internal, external, and contextual. I believe that the formative evaluator will have to collect all three types of information and that he will have to search for links between context, the product, and its effects. Let me illustrate this last point by again comparing a formative evaluator with a summative evaluator.

A summative evaluator might well make use of the three classes of information described in this paper. For instance, the summative evaluator might determine that the physical characteristics of the product are compatible with his own needs. Or he might determine that the contextual characteristics listed for the materials are appropriate for his particular situation. Or he might utilize information that the materials have had demonstrated effects on students and teachers. The summative evaluator might make his decision about whether or not to purchase these materials upon the basis of any one of these considerations or upon all of these considerations simultaneously.

The formative evaluator will probably make somewhat different use of these data, however. He will seek to determine relationships between contextual, internal, and external information. When the formative evaluator collects information that his materials have failed to reach certain objectives, he must ask the question, "Why?" at this point the formative evaluator begins to function very much like a researcher. He hypothesizes various possible causes of an observed result and tests them in a more or less systematic manner. For instance, suppose a particular

textbook teaches concepts A, B, and C very well but fails to teach concepts D and E and provides only partial understanding of concept F. What should be done to correct this situation? The evaluator might hypothesize that the readability level of the text at those portions which teach concepts D, E, and F exceed the reading level of target students, or that the sequencing of concepts should be F, D, E, not D, E, F, or that instructional aids that were presumed to be available were not, or that teachers do not possess sufficient subject matter competence to employ these materials, and so on. The formative evaluator will seek out relevant information of this sort so that the appropriate adjustments can be made in the intended contexts, in the product itself, or in the intended outcomes.

Consider another example. Suppose it is determined that students for whom the materials are intended have a rather low opinion of the materials and express a negative reaction toward their use. A summative evaluator might well consider such information sufficient for rejecting the materials. The developer of the materials needs more information. He needs to know <u>why</u> students have negative attitudes. He needs to know whether students who are actually using the materials have such negative attitudes, and, if so, how he might alter his materials or the intended contexts to reduce these negative appraisals. Again the formative evaluator will be inquiring into the "why of outcomes" (Hastings, 1966), into the relationship between context, internal, and external characteristics of the instructional materials.

One final point needs to be made before closing this discussion. Different types of information and different methodologies of formative evaluation are most useful at different stages in the development of materials. For instance, the author appraisal and oral problem-solving approaches mentioned earlier would be very appropriate and potentially

very useful in the early stages of development. But these same sources should be relied on less and less as successive revisions are performed. For an author to change his materials (in any substantial manner) on the basis of hindsight after considerable external information has substanciated the effectiveness of the materials, would be foolish. Similarly, field testing with large groups of students at a very early age in the development of a product would be wasteful of both time and effort. Much more useful information is likely to be gained from a single student who is primed to think aloud.

Anderson, R.C. The Comparative Field Experiment: An Illustration from High School Biology. Proceeding of the 1968 Invitational Conference of Testing Problems, Princeton, N.J.: Educational Testing Service, 1969, Pp. 3-30.

Berelson, B. "Content Analysis", in Handbook of Social Psychology, Vol. I., 1954, Pp. 458-518. Cambridge, Mass.: Addison-Wesley.

Bloom, B.S. and Broder, L.J. Problem-Solving Process of College Students. 1950, Chicago: University of Chicago Press.

Cronbach, L.J. (Ed.) Text Materials in Modern Education, 1955, University of Illinois Press, Urbana, Illinois.

Cronbach, L.J. Course Improvement Through Evaluation. Teachers College Record, 1963, 64, 675-683.

Eash, M.J. Assessing Curriculum Materials: A Preliminary Instrument. Educational Product Report, 1969, 2, 18-24.

Easley, J.A. A Project to Develop and Evaluate a Computerized System for Instructional Response Analysis. Computer-based Education Research Laboratory. University of Illinois, Urbana, Ill., 1967, (mimeo).

Easley, J.A., Jr., Jenkins, E.S., and Ashenfelter, J.W. A Scheme for the Analysis of Elementary Science Materials. The EPIE Forum, I, No. 3; 16-21, 1967.

Edmonson, J.B., et al. The Textbook in American Education. Thirtieth yearbook of the National Society for the Study of Education, Part II. Bloomington, Illinois, Public School Company, 1931.

Gagne, R.M. The Conditions of Learning, 1970, Holt, Rinehart, and Winston, New York.

Ginser, R. Instructional Technology and the Measurement of Learning Outcomes: Some Questions. American Psychologist, 1963, 18, 519-521.

Guba, E.G. The Failure of Educational Evaluation. Educational Technology. 1969, 9, No. 5, 29-39.

Hastings, J.T. Curriculum Evaluation: The Whys of the Outcomes. Journal of Educational Measurement, 1966, 3, 27-32.

Hoban, C.F. Focus on Learning, 1942, Washington, D.C., American Council on Education.

Kerlinger, F. Foundations of Behavioral Research, 1964, Holt, Rinehart, and Winston, New York, Ch. 30.

Kibler, R.J., Barker, L.L., and Miles, D.T. Behavioral Objectives and Instruction. Allyn and Bacon, Inc., Boston, 1970.

Klaus, D.J. An Analysis of Programming Techniques. In Glaser, R. (Ed.) Teaching Machines and Programmed Learning, II; Data and Directions. National Education Association, Washington, D.C., 1965.

Kracauer, Sigfried, Theory of Film, Oxford University Press, New York, 1960.

Krathwohl, D.R. The Taxonomy of Educational Objectives--Use of the Cognitive and Affective Domains. In C.M. Lindvall (Ed.) Defining Educational Objectives, 1964, University of Pittsburg Press, Pittsburg, Pennsylvania.

Krathwohl, D.R. Stating Objectives Appropriately for Program, for Curriculum and for Instructional Materials Development, Journal of Teacher Education, 12, 83-92, 1965.

Lumsdaine, A.A. Assessing the Effectiveness of Instructional Programs. In Glaser, (Ed.) Teaching Machines and Programmed Learning, II, Data and Directions. 1905, National Education Association, Washington D.C.

Mager, R.F. Preparing Instructional Objectives, Palo Alto, California, Fearon Publishers, 1962.

Markle, S.M. Good Frames and Bad. (2nd Ed.) New York, 1969, John Wiley and Sons.

Metfessel, N.S. and Michael, W.B. A Paradigm Involving Multiple Criterion Measures for the Evaluation of the Effectiveness of School Programs. Educational and Psychological Measurement, 1967, 27, 931-943.

Rosenshine, B. Evaluation of Instruction. Review of Educational Research, 1970, 40, 279-300.

Scott, R.O. and Yelon, S.L. The Student as a Co-Author--The First STep in Formative Evaluation. Educational Technology, 1969, 9, No. 10, 76-78.

Scriven, M. The Methodology of Evaluation. In R.E. Stake (Ed.) Perspectives of Curriculum Evaluation, AERA monograph series on Curriculum Evaluation, No. 1, 1967, Chicago, Ill., Rand McNally.

Sjogran, D.D. Measurement Techniques in Evaluation. Review of Educational Research, 1970, 40, 301-320.

Spottiswoode, Raymond. Film and Its Techniques, University of California Press, Berkeley and Los Angeles, California, 1958.

Stake, R.E. The Countenance of Educational Evaluation. Teacher College Record, 1967, 68, 523-540.

Stake, R.E. Generalizability of Program Evaluation: The Needs for Limits. Educational Product Report, 1969, 2, 39-40.

Stake, R.E. Objectives, Priorities, and other Judgment Data. Review of Educational Research, 1970, 40, 181-212.

Stevens, W.W., Jr. and Morressett, I. A System for Analyzing Social Science Curricula. The EPIE Forum, 1967, 1, Nos. 4 and 5, Pp. 10-15.

Stufflebeam, D. Evaluation as Enlightenment for Decision Making. Columbus, Ohio Evaluation Center, Ohio State University, 1968.

Tyler, L.T. Recommendations for Evaluating Curriculum and Instructional Materials. Paper presented at the annual meeting of the American Educational Research Association, Minneapolis, Minnesota, February, 1970.

Westbury, I. Curriculum Evaluation. Review of Educational Research, 1970, 40, 239-260.